15. Johnson LC, Letzel HW (1984) Measuring passive smoking: methods, problems, and perspectives. Prev Med 13:705-716
16. Koo LC, Ho JH-C, Saw D (1984) Is passive smoking an added risk factor for lung cancer in Chinese women? J Exp Clin Cancer Res 3:277-283
17. Lee PN, Chamberlain J, Alderson MR (1986) Relationship of passive smoking to risk of lung cancer and other smoking-associated diseases. Br J Cancer 54:97-105
18. Lehnert G (Chairman), Garfinkel L, Hirayama T, Schmahl D, Überla K, Wynder EL, Lee P (1984) Roundtable discussion. Prev Med 13:730-746
19. Lubin JH, Blot WJ (1984) Assessment of lung cancer risk factors by histologic category. JNCI 73:383-389
20. MacLennan R, Da Costa J, Day NE, Law CH, Ng YK Shanmugaratnam K (1977) Risk factors for lung cancer in Singapore Chinese, a population with high female incidence rates. Int J Cancer 20:854-860
21. Muramatsu M, Umemura S, Fukui J, Arai T, Kira S (1987) Estimation of personal exposure to ambient nicotine in daily environment. Int Arch Occup Environ Health 59:545-550
22. National Research Council, Committee on Passive Smoking, Board on Environmental Studies (1986) Environmental tobacco smoke. Measuring exposures and assessing health effects. National Academy Press, Washington, DC, pp 224-227
23. Pershagen G, Hrubec Z, Svensson C (1987) Passive smoking and lung cancer in Swedish women. Am J Epid 125:17-24
24. Rosner B (1982) Fundamentals of Biostatistics. Duxbury Press, Boston, MA, p 176
25. Sellers TA, Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE, Rothschild H (1987) Increased familial risk for non-lung cancer among relatives of lung cancer patients. Am J Epid 126:237-246
26. Trichopoulos D, Kalandidi A, Sparros L, MacMahon B (1981) Lung cancer and passive smoking. Intl J Cancer 27:1-4
27. U.S. Department of Health and Human Services, Public Health Service (1982) The Health Consequences of Smoking: Cancer: a report of the Surgeon General. DHHS Publication No. (PHS)82-50179, p 243
28. U.S. Department of Health and Human Services. Public Health Service (1986) The health consequences of involuntary smoking. Report of the Surgeon General. DHHS Publication No. (CDC) 87-8398, pp 66-102
29. U.S. Department of Health, Education, and Welfare, Public Health Service (1964) Report of the Advisory Committee to the Surgeon General of the Public Health Service. Smoking and health. PHS Publication No. 1103, pp 174-175
30. Wu AH, Henderson BE, Pike MC, Yu MC (1985) Smoking and other risk factors for lung cancer in women. JNCI 74:747-751

Passive Smoking and Lung Cancer:
A Reanalysis of Hirayama's Data

K. Überla and W. Ahlborn

The statistical association between environmental tobacco smoke and lung cancer is controversial. The Hirayama Study seems to provide sound epidemiological evidence supporting this hypothesis. In a recent paper [6] I have analyzed the published studies. Regarding the Hirayama study the following facts have to be kept in mind:

- The study was not designed to test the hypothesis, whether passive smoking is associated with lung cancer or not. It can therefore only generate this hypothesis, not prove it.
- The cohort was not representative for the population of Japan. A selection bias is possible.
- The exposure indicator - the fact of being married to a man who smokes - is not reliable, not valid and not specific.
- The event indicator - dying on lung cancer as noted on death certificates - is neither reliable nor valid.
- Various confounding factors - for instance exposure at the working place, indoor air pollution, overall air pollution, type of medical care - were not accounted for.
- Bias in registering the fact, that a woman is a nonsmoker, was not controlled. Resulting differential misclassifications of the cases, who were smokers and had to be excluded, have not been considered.
- Almost nothing is known about the 200 cases. No case reports are available, autopsy and histology are available in only 11.5%.

The core of the information, on which the results of this study rely, is

1) that during 1965 200 women in Japan told an interviewer on a single occasion that they were - during that time - nonsmokers and their husbands told that they were smokers, which might have been different before and afterwards and
2) that their death certificates subsequently contained the diagnosis lung cancer, which might have been erroneous.

Such sparse information does not seem to be convincing.

In our paper we consider three questions:

1) What is the relative risk when one removes the selection bias regarding age of women in the Hirayama cohort?
2) What is the relative risk for women married to men with different occupations, when one removes the selection bias regarding age of men?
3) What is the relative risk when additionally some differential misclassification is assumed?

Material and Methods

We start from Tables 1, 2, and 3 of Hirayama 1984 [4]. These tables contain the most detailed published data. In order to check our program, we reproduced some of the reported relative risk estimates with good accuracy.

There are marked differences between the Hirayama cohort and the female age distribution over 40 in the population of Japan 1965. Women 50–59 are overrepresented, women older than 70 are severely underrepresented. In this age group only a single case from 12 was observed. The investigated cohort certainly has a severe selection bias by age, which needs no statistical test. This is likely due to the fact, that the smoking behaviour was not known in the elderly or that the husbands of older women have died. Since it takes 20 years and more from exposure to lung cancer, older women surely are relevant and should not be excluded. The majority of lung cancer cases occur in older age groups, in Germany more than 67% in women over 65 years.

In order to answer the question what the relative risk is when the age selection bias is removed, we adjusted the data to the age distribution of the female population of Japan.

Table 1. Differences between Hirayama cohort and the female age distribution over 40 in the population of Japan 1965*

Age group	Percent female	
	Japan population	Hirayama cohort
40–49	39	42
50–59	30	35
60–69	19	22
70 +	12	11
	100	100

* Population Census 1965. Statistical survey of economy of Japan; 1967. Ministry of Foreign Affairs of Japan.

Table 2. Smoking habit of husband by age of wife.* Original data

Wives age	Husbands smoking habit							
	Non		1-19		20 +		Total	
40–49	4	7,918	21	17,492	21	12,615	46	38,025
50–59	14	7,635	46	15,640	31	8,814	91	32,089
60–69	16	6,170	31	10,381	10	3,793	57	20,344
70 +	3	172	1	671	2	239	6	1,082
Total	37	21,895	99	44,184	64	25,461	200	91,540

* Table 2 of Hirayama 1984.

Table 3. Smoking habit of husband by age of wife*. Removed selection bias: Data adjusted to the age distribution of women in the population

Wives age	Husbands smoking habit							
	Non		1-19		20 +		Total	
40–49	3.91	7,748.8	19.12	15,927.8	20.02	12,024.0	43.05	35,700.6
50–59	12.49	6,813.7	38.20	12,987.1	26.95	7,661.2	77.64	27,462.0
60–69	14.25	5,496.6	25.70	8,604.9	8.68	3,291.1	48.63	17,392.6
70 +	32.02	1,835.9	9.93	6,664.2	20.79	2,484.7	62.74	10,984.8
Total	62.67	21,895	92.95	44,184	76.44	25,461	232.06	91,540

* Table 2 of Hirayama 1984.

The technique of iterative proportional fitting of a contingency table to given marginals as described by Bishop et al. [1] or by Hartung et al. [3] was used. This technique keeps the risks constant as observed in every cell and changes the marginals and the cell counts according to the given age distribution of the population. Iterative proportional fitting of contingency tables to given marginals is a well known technique in multivariate statistics and can be applied here without changing the observed interrelations between smoking habit, occupation, and lung cancer. From the fitted or adjusted tables the risk ratios are calculated in the usual way. Such risk ratios based on data with removed age selection bias are the correct ones and should be used.

One has to require that there should be no selection bias by age and the cases should be included as they would have occured in the population. Otherwise statistical tests and p-values are not very meaningful. Table 2 shows the original data by age of wife. The cells contain the number of lung cancer cases and those under risk as published by Hirayama. The 1–19 group includes ex-smokers in this and the following tables. 200 cases out of 91,540 women were observed. Iterative proportional fitting to the female age distribution of the population leaves the hatched numbers constant. The others are adjusted using a right hand marginal which is made proportional to the age distribution of the population.

Results

Table 3 gives the results of iterative proportional fitting to the female age distribution of the population. It contains the numbers of those under risk and of lung cancer deaths as they would have been observed, if Hirayama had not excluded or preferred certain age groups. The age selection bias is removed. The risks in the individual cells are still the same as those observed by Hirayama. Also the structure of the common distribution regarding age, smoking habit and lung cancer is unchanged. Hirayama would have totally observed 232 cases instead of 200, with the corresponding numbers in the individual cells, had he included all women as they live in the population. This table is the best available starting point for age-standardized risk ratio calculations. It was not used so far.

Table 4. Relative risk by age of women*

| | Husbands smoking habit | | |
	Non	1-19	20 +
$\hat{R}R$	1.00	1.37	1.56
IL_{90}		1.00	1.11
MH-CHI		1.51	2.27
$P_{one\ tailed}$		0.065	0.012**
$\hat{R}R$	1.00	0.77	1.06
IL_{90}		0.59	0.80
MH-CHI		2.19	0.27
$P_{one\ tailed}$		0.014***	0.395

Upper part: standardized by age of women only.
Lower part: age selection bias removed and standardized by age of women.
$\hat{R}R$: Weighted point estimate of rate ratio.
IL_{90}: Lower 90-percent confidence interval.
* Calculated from Table 2 of Hirayama 1984.
** "Significant" in positive direction.
*** "Significant" in negative direction.

In the upper part of Table 4 you find the risk ratios standardized by age only, as done by Hirayama. The lower part are the risk ratios after removing the age selection bias. In the upper part the weighted point estimate of the rate ratio is 1.56 in the 20+-group and is technically "significant". IL_{90} designates the lower point of the 90-percent confidence interval in this and the following tables, as it was used by Hirayama. This risk increase disappears completely when one removes the selection bias by age. In the 20+-group the rate ratio is 1.06, hardly a relevant risk increase. In the group of 1-19 cigarettes per day it is 0.77 which is a technically significant risk decrease. The adjusted rate ratio, considering all those exposed in one group versus those not exposed is 0.901 with a confidence interval including unity. If Hirayama had observed the cases as they occur in the female population without selection bias by age, he would have observed no risk increase, but a risk decrease. This is the main result of our reanalysis, which corresponds well with the result of the prospective American cohort study as published by Garfinkel [2].

We now consider two occupations, farmers and industry workers. From the upper part of Table 5 one can see that the relative risk for wives of farmers seems substantial, when one standardizes by age of men only. The point estimates of the rate ratios are 1.48 and 1.63 respectively. This was observed earlier and had no adequate explanation. If one removes the selection bias by age and adjusts to the male age distribution of Japan – the numbers in the lower part of Table 5 – the rate ratios are 0.85 and 0.82, not different from unity. This seems more plausible.

Considering the wives of industry workers only, in the upper part of Table 6, the point estimates of the rate ratios are 1.77 and 2.27, standardized by age of men, being not significant. Removing the age selection bias – in the lower part of Table 6 – there is a remarkable risk increase to 4.60 and 6.90, which is significant. However, there are only

Table 5. Relative risk: wives of farmers only*

| | Husbands smoking habit | | |
	Non	1-19	20 +
$\hat{R}R$	1.00	1.48	1.63
IL_{90}		0.97	1.01
MH-CHI		1.48	1.92
$P_{one\ tailed}$		0.069	0.027
$\hat{R}R$	1.00	0.85	0.82
IL_{90}		0.59	0.53
MH-CHI		0.42	0.53
$P_{one\ tailed}$		0.337	0.296

Upper part: standardized by age of men only.
Lower part: age selection bias removed and standardized by age of men.
$\hat{R}R$: Weighted point estimate of rate ratio.
IL_{90}: Lower 90-percent confidence interval.
* Calculated from Table 3 of Hirayama 1984.

Table 6. Relative risk: wives of industry workers only*

| | Husbands smoking habit | | |
	Non	1-19	20 +
$\hat{R}R$	1.00	1.77	2.27
IL_{90}		0.70	0.84
MH-CHI		0.73	0.81
$P_{one\ tailed}$		0.232	0.208
$\hat{R}R$	1.00	4.60	6.90
IL_{90}		1.71	2.45
MH-CHI		2.50	2.78
$P_{one\ tailed}$		0.006	0.003

Upper part: standardized by age of men only.
Lower part: age selection bias removed and standardized by age of men.
$\hat{R}R$: Weighted point estimate of rate ratio.
IL_{90}: Lower 90-percent confidence interval.
* Calculated from Table 3 of Hirayama 1984.

9 lung cancer deaths in the 20+-group and only 3 in women 70 years and older, which are small numbers, but these are numbers observed and used by Hirayama and his risk structure is unchanged. Thus only in the subgroup of women married to industry workers there is a risk increase, in all other occupations there is no risk increase. Omitting industry

Table 7. Relative risk: assumed differential misclassifications*

Number of cases assumed misclassified and removed from exposed groups		Husbands smoking habit		
		Non	1-19	20 +
n = 10 = 5%	$\hat{R}R$	1.00	0.74	1.00
	$P_{one\ tailed}$		0.006	0.469
n = 20 = 10%	$\hat{R}R$	1.00	0.70	0.93
	$P_{one\ tailed}$		0.003	0.383
n = 30 = 15%	$\hat{R}R$	1.00	0.66	0.85
	$P_{one\ tailed}$		0.001	0.238

Age selection bias removed and standardized by age of women.
$\hat{R}R$: Weighted point estimate of rate ratio.
* Calculated from Table 2 of Hirayama 1984.

Table 8. Reanalysis of Hirayama's data: summary of relative risk

		Husbands smoking habit		
		Non	1-19	20 +
Age selection bias removed and age-standardized (women)	$\hat{R}R$	1.00	0.77	1.06
	$P_{one\ tailed}$		0.014	0.395
Without industry workers, age selection bias removed and age-standardized (men)	$\hat{R}R$	1.00	0.90	0.89
	$P_{one\ tailed}$		0.394	0.179
10 cases assumed misclassified, age selection bias removed and age-standardized (women)	$\hat{R}R$	1.00	0.74	1.00
	$P_{one\ tailed}$		0.006	0.469

$\hat{R}R$: Weighted point estimate of rate ratio.

workers, the point estimates of the rate ratios are 0.90 and 0.89, not significantly different from unity. These findings are consistent with the assumption of confounding factors in women married to industry workers, who might be exposed to other environmental hazards. Our calculations show that by removing selection bias by age, one can explain hitherto implausible results.

Active smoking is correlated among married couples. In a society in which female smokers were very rare in 1965, more women married to smokers will declare themselves nonsmokers than the other way round. One has therefore to considerr biased or differential misclassification. There are likely more women with lung cancer, who have been misclassified as nonsmokers and have to be removed from the cohort, than the other way round.

We made some moderate assumptions regarding differential misclassification, as shown in Table 7. In order to examine how sensitive the relative risk is we removed 10, 20, and 30 cases from the exposed groups - corresponding to 5, 10, and 15 percent.

Assuming 30 misclassified cases - 15 percent, a percentage which has been observed in the literature [5] - the rate ratios are 0.66 and 0.85. In the group 1-19 cigarettes per day all the risk estimators are significantly smaller than unity. Our personal opinion is that 10 differential misclassified cases from 200, who have to be omitted, are a fair number. The corresponding weighted point estimates of the rate ratio are 0.74 and 1.00. These risk estimates are as reasonable as other risk estimates calculated from the Hirayama data. They indicate - if anything - a risk decrease, not a risk increase.

Discussion

Reanalyses of data, which have been collected by others are not easy. This is because information is not completely available, because information might be misinterpreted or because one has to take another view in order to come closer to the acceptable truth. Our calculations do not diminish the great value and impact the Hirayama study had on the epidemiology of passive smoking. They show however, that reasonable alternative views

on the same data are possible, which lead to opposite conclusions. Our findings are in contrast to Hirayama's thesis that - based on his data - there is a substantial statistical association between passive smoking and lung cancer.

As long as there is no other independent and sound epidemiological evidence, it should be left to the individual scientist which analysis of the same data he thinks is more appropriate. We do not hold that our view is the only correct one. We do hold however, that the risk ratios calculated by us, removing age selection bias, are as valid as other risk estimates. To our opinion they are more appropriate, since they go back to the population and not to a selected sample. Even when one would take another marginal, for instance the age distribution of wives still married to living men - which was not available - the effect would be considerable. Our risk estimates are a consequence of the data published by Hirayama and cannot be rejected from the study data, as they are published so far.

To summarize (Table 8): Removing the age selection bias in the Hirayama study one gets a relative risk of 1.06 in the group of women married to men with more than 20 cigarettes per day. In the group of women married to men with 1-19 cigarettes per day the relative risk is 0.77, a technically "significant" risk decrease. If Hirayama could have observed the lung cancer cases as they occur in the female population, he would have observed no risk increase, but a risk decrease to around 0.90, considering those exposed versus those not exposed. This fact deserves attention.

If one omits the wives married to industry workers because of possible confounding factors in this group, the relative risk is 0.90 and 0.89 respectively. This is of the same size order and smaller than unity. Here we could adjust and standardize by occupation and age of men only, which is not as appropriate as by the age of women.

If one assumes that 10 cases are differentially misclassified and removes them from the exposed groups, the risk estimates are 0.74 and 1.00, respectively. Our findings demonstrate how sensitive the data of this study are and how weak the evidence for a statistical association between passive smoking and lung cancer might be. In view of these and other facts some of which we mentioned in the introduction, the null hypothesis might be true as well and seems to be consistent with the Hirayama data in the same way as the alternative hypothesis.

Unrepresentativeness

Although it is clear that the combined study population in Table 1 is not fully representative of non-smokers, this does not appear to be a serious issue, since studies have been carried out in the US, UK, Greece, Sweden, Japan and Hong Kong, and Wald et al. (1986) found no evidence of significant heterogeneity of relative risk estimates.

Sample Size

While none of the studies considered in Table 1 concern particularly large numbers of non-smokers with lung cancer, chance can hardly be the total explanation of the association between passive smoking and lung cancer since the overall relative risk estimate quoted by Wald et al. (1986) of 1.35 is quite highly statistically significant, with 95% confidence limits of 1.19-1.54.

Confounding

Not all the studies considered have taken into account the possible confounding effect of factors known or suspected to be related to lung cancer, such as occupation or nutrition. Although it would perhaps be expected that non-smokers married to smokers might to some extent share the tendency of smokers to work in dirtier jobs, standardisation for occupation, or indeed any confounding factor, has never been found in practice to explain any material part of the association between lung cancer and passive smoking. It does not seem likely that failure to take confounding factors into account has materially affected the issue.

Inappropriate Choice of Controls

General scientific principles demand that like should be compared with like as far as possible. In a number of studies, there were clear exceptions to this. One example is the study of Trichopoulos et al. (1981, 1984) in which controls came from a different hospital. This may cause bias if patients came from different catchment areas with different smoking characteristics. Another example is the recently reported study of Humble et al. (1987), in which virtually all the interviews with controls were conducted directly while much of the data for cases came from surrogates. While inappropriate choice of controls may have materially biassed a few studies, it does not seem very likely, however, that it is a major explanation for the huge discrepancy.

Inaccuracy of Disease Classification

It is well known that diagnosis of lung cancer is imperfect and studies such as those by Garfinkel et al. (1985) which took pains to check and review all available evidence are to be preferred to those that did not do so. However, random misclassification of diagnosis would be expected to reduce the observed association between lung cancer and passive smoking, not increase it, and differential misclassification of diagnosis does not seem very likely, inasmuch as the doctor making the diagnosis is likely to have been blind to the

patient's ETS exposure in most cases, and not to have been affected by it even if he was, given most diagnoses were made before the first study on the issue was published in 1981.

Non-reporting Bias

A problem in combining results from various studies to come to an overall assessment of the evidence, by so-called "meta-analysis", is the possibility that the studies being combined are not representative of all those that have been carried out. In particular, overall estimates of relative risk may be biassed upward if a scientist is less likely to submit for publication, or a journal is less likely to publish, studies which show no significant relationship of disease to the factor of interest or a significant trend in the direction opposite to that expected in advance. Convincing evidence of bias resulting from this in the context of randomised controlled trials has recently been collected by Chalmers et al. (1987) and the problem may generally be greater in epidemiological studies. One can easily imagine an investigator running a range of statistical analyses, finding a few significant associations of interest, and then publishing papers on those, ignoring the non-significant relationships. One can also imagine journals not being too keen to given space to a paper on a new null association and one must inevitably wonder whether the reason the first two studies published on passive smoking and lung cancer (Hirayama 1981; Trichopoulos 1981) found a significant association was because first published papers on any association tend to be positive. Indeed, there seems a case for carrying out meta-analysis giving most weight to studies showing no association and least weight to studies published first.

Although it is obviously important to conduct research into the problems of non-reporting bias in epidemiological studies, it is difficult to claim that it is the full explanation of the overall association between passive smoking and lung cancer. The reasons for this view are two-fold. Firstly, the overall association remains significant (though the relative risk estimate reduces to 1.20), even after eliminating the Hirayama and Trichopoulos results. Secondly, the fact that lung cancer and passive smoking has been a very "hot" issue in recent years suggests researchers should now be able to publish results from studies showing no association between passive smoking and lung cancer risk.

Lack of Objective Measure of ETS Exposure

A limitation of all the published epidemiological evidence is lack of objective measurement of exposure to ETS. Subjects are classified mainly by whether or not they are married to a smoker and occasionally by reported degree of exposure outside the home, but there are no data available either on ambient levels of tobacco smoke constituents at home or at work or on levels in body fluids such as blood, urine or saliva. While (e.g. see Table 3) it can be shown that marriage to a smoker is indeed associated with increased levels of cotinine, the relatively crude method used for determining exposure leads to possibilities of bias in case-control studies where knowledge of disease may consciously or subconsciously affect reporting of ETS exposure. While this may have caused upward bias of the reported relative risk in some case-control studies, it can hardly explain the whole association, since it would not be expected to cause upward bias in prospective studies and the association seems as strong in prospective as in case-control studies.

Table 5. Hypothetical example of bias due to misclassification of 5% of smoking subjects as non-smokers

Smoking habit*		Assumed		Observed			
Subject	Spouse	N	Risk	N	Risk	Passive effect	Active effect
NS	NS	60	1	60 + 2 = 62	1.61	1	
NS	S	40	1	40 + 3 = 43	2.33	1.44	
NS	Total	100	1	100 + 5 = 105	1.90		1
S	NS	40	20	40 − 2 = 38	20		
S	S	60	20	60 − 3 = 57	20		
NS	Total	100	20	100 − 5 = 95	10		10.5

Assumed concordance $= (60 \times 60)/(40 \times 40) = 2.25$
Observed concordance $= (62 \times 57)/(43 \times 38) = 2.16$

* NS, non-smoker; S, smoker

Lack of Objective Measure of Active Smoking Status

Although considered last, this appears to be the most serious problem affecting the epidemiological evidence on passive smoking and lung cancer. As will be shown in the next section, completely erroneous conclusions can be reached when the "non-smokers" being studied actually include a small proportion of misclassified true smokers.

Misclassification of Active Smoking Habits as a Major Source of Bias

As shown in Table 5, misclassification of a small proportion of smokers as non-smokers, coupled with a tendency for smokers to be married to smokers ("concordance") can create an apparent positive effect of passive smoking when no actual effect exists. It also leads to an underestimation of the active smoking effect and of the concordance. The passive smoking bias depends critically on the assumed relative risk for active smoking, the degree of concordance and on the level of misclassification of subject smokers as non-smokers. This source of bias will also produce an artificial dose response relationship when the "non-smoking" subjects are divided according to the amount smoked by the spouse. It can be shown (Lee 1988b) that misclassification of non-smoking subjects as smokers and of smoking spouses as non-smokers causes a degree of bias that is minor compared with that resulting from misclassification of smoking subjects as non-smokers.

In an attempt to determine the extent to which smokers misreported their smoking habits and to which smokers tend to be married to smokers, Lee (1987) carried out three separate studies. In the first study, which concerned accuracy of reported current habits, 1775 British subjects were asked about their smoking habits and use of other nicotine products in a non-health context likely to minimize underreporting of smoking. They were then (with no prior warning) asked to provide saliva for cotinine analysis and 1537 agreed to do so. As shown in Table 3 there was in general a very marked difference between the cotinine levels of tobacco users and non-users. Using 30 ng/ml as a cut-off,

1.1% of self-reported non-users could be classified as occasional users, with levels of to 100 ng/ml, while 1.4% could be classified regular users, with levels above 100 ng/ml.

The second study, which aimed at obtaining information, on accuracy of past smoking habits, followed up in 1985 540 subjects previously interviewed in 1980 about their smoking habits. Ten% claiming on one occasion never to have smoked made inconsistent statements on the other occasion, with inconsistent smokers being more often men, old, smokers of fewer cigarettes and long term ex-smokers.

The third study, which aimed at obtaining information on smoking habit concordance, involved 8857 subjects aged 16+ interviewed regarding their own smoking habits and that of their spouse. The concordance ratio, 3.55 in men and 3.07 in women, was found to be rather greater than that assumed in the example in Table 5. Concordance rose with amount smoked. Thus, the chance of having a spouse who was a manufactured cigarette smoker was 22% for subjects who reported no such smoking, and 45%, 52% and 59% respectively for subjects who reported smoking 1-17, 18-22 and 23+ manufactured cigarettes per day.

From the data obtained, Lee (1987) concluded that misclassification could bias relative risk estimates in relation to passive smoking upwards by a factor of 1.31 in men and 1.24 in women, not significantly different from the pooled estimate of risk in relation to passive smoke exposure.

At about the same time as Lee presented his findings, Wald et al. (1986) used similar techniques to estimate the bias from misclassification, but based on a number of smaller, and less representative studies, not specifically designed for the purpose. They estimated this misclassification would have less effect, reducing the pooled estimate of risk only from 1.35 to 1.30, i.e. it had only caused upward bias by a factor of 1.04.

Examination of the detail of how the estimates of bias of Wald et al (1986) and of Lee (1987) were arrived at reveals three reasons for the difference. The first was that Wald et al. (1986) used an assumed relative risk of 8 for the effect of active smoking observed in women whereas Lee (1987) used 10. The second was that the calculations of bias by Wald et al. (1986) were mathematically inaccurate, due to confusion between true relative risks in relation to active smoking and those observed (which are affected by misclassification). These are less important than the third reason, which is that Lee et al. (1987) found that 1.4% (10/808) self-reported non-smokers were current regular smokers, whereas Wald et al. (1986) only found 0.14% (1/705) such cases.

In an attempt to reconcile this difference, I have recently conducted a detailed literature review of the evidence on misclassification of smoking habits, which will be published as a book early in 1988.

Despite the various study designs and populations involved a number of clear conclusions were reached:

(1) Even in circumstances that are apparently similar quite a wide variation in the extent of misclassification can be found.

(2) The proportion of "non-smokers" subsequently found actually to be smokers is markedly higher in smoking cessation studies than in studies where the respondent is under no special pressure not to smoke.

(3) The proportion of "non-smokers" subsequently found actually to be smokers is also markedly higher in lung cancer patients than in the general population. This is not surprising in view of the overall *a priori* expectation that a lung cancer patient actually is a smoker.

(4) Studies of "non-smokers" without lung cancer and under no special pressure not to smoke suggest that around 4% are likely actually to be current smokers. While not all

studies provide information on the extent to which such misclassified smokers smoke, and those that do indicate many of them are occasional smokers, it seems that 1 to 2% of self-reported non-smokers are regular smokers.

(5) In addition to these misclassified current smokers there are a somewhat larger number of ex-smokers misclassified as never smokers. Available information suggests that these tend to have smoked less and a longer time ago than average ex-smokers.

(6) None of the studies have investigated whether the extent to which smokers deny smoking depends on whether their spouse happens to smoke, which is of theoretical importance as it could materially affect estimates of bias.

(7) There is even now virtually no information on the extent to which smoking habits might be misclassified in Japan and Greece, from whence came the early epidemiological evidence on passive smoking and lung cancer. The only study in Japan, by Akiba et al. (1986), provided data suggestive of substantial misclassification of smoking habits. Here, of 187 men who reported not smoking in 1964–68, as many as 96 (51%) reported in 1982 that they had smoked.

While there is an obvious need for further research on misclassification of smoking habits the results of the review[1] suggested strongly that Wald et al. (1986) had seriously underestimated its importance.

Overall Conclusion

It has clearly been shown that there is a huge discrepancy between the relative doses of smoke constituents to which passive and active smokers are exposed and the much larger relative effect claimed from epidemiological evidence. The most likely explanation for this discrepancy seems to be a persistent bias affecting all the studies due to a small proportion of smokers being wrongly classified as non-smokers in the epidemiological studies.

References

Adlkofer F, Scherer G, von Hees U (1985) Passive smoking. N Engl J Med 312:719–720

Akiba S, Kato H, Blot WJ (1986) Passive smoking and lung cancer among Japanese women. Cancer Research 46:4804–4807

Alderson MR (1983) An introduction to epidemiology. Macmillan

Alderson MR, Lee PN, Wang R (1985) Risks of lung cancer, chronic bronchitis, ischaemic heart disease and stroke in relation to type of cigarette smoked. J Epidemiology and Community Health 39:286–293

Arundel A, Irwin T, Sterling TK, Weinkam J (1986) Nonsmoker lung cancer risks from exposure to tobacco smoke particulates. Stated to be available on request in: Arundel A, Irwin T, Sterling T. Nonsmoker lung cancer risks from tobacco smoke exposure: An evaluation of repace and lowrey's phenomenological model. J Environmental Science and Health Part C 4:93–118

Buffler PA, Pickle LW, Mason TJ, Contant C (1984) The causes of lung cancer in Texas. In: Lung cancer causes and prevention. Chemie International Inc, pp 83–99

[1] Supported by a recent large study by Coultas et al. (American Review of Respiratory Diseases, 1987, 136, 305–309) who found 63 subjects with salivary cotinine >20 ng/ml among a sample of 1,360 New Mexicans reporting they were not current smokers.

Chalmers TC, Levin H, Sachs HS, Reitman D, Berrier J, Nagalingam R (1987) Meta-analysis of clinical trials as a scientific discipline. I: Control of bias and comparison with large co-operative trials. Statistics in Medicine 6:315–325

Chan WC, Fung SC (1982) Lung cancer in non-smokers in Hong Kong. Cancer campaign. In: Cancer epidemiology. Fischer, vol 6, 199–202

Correa P, Pickle LW, Fontham E, Lin Y, Haenszel W (1983) Passive smoking and lung cancer. Lancet 2:595–597

Dalager NA, Pickle LW, Mason TJ, et al (1986) Relation of passive smoking to lung cancer. Cancer Research 46:4808–4811

Doll R (1986) Lung cancer observed and expected changes in incidence from active and passive smoking. XIV UICC Conference, Budapest

Doll R, Peto R (1978) Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and lifelong non-smokers. J Epidemiology and Community Health 32:303–313

Garfinkel L (1981) Time trends in lung cancer mortality among non-smokers and a note on passive smoking. J National Cancer Institute 66:1061–1066

Garfinkel L, Auerbach O, Joubert L (1985) Involuntary smoking and lung cancer: A case-control study. J National Cancer Institute 75:463–469

Gillis CR, Hole DJ, Hawthorne V, Boyle P (1984) The effect of environmental tobacco smoke in two urban communities in the west of Scotland. Eur J Resp Dis [Suppl 133] 65:121–126

Hiller FC, Mazumder MK, Wilson JD, McLeod PC, Bone RC (1982a) Human respiratory tract deposition using multimodal aerosols. J Aerosol Science 13:337–343

Hiller FC, McCusker KT, Mazumder MK, Wilson JD, Bone RC (1982b) Deposition of sidestream cigarette smoke in the human respiratory tract. American Review of Respiratory Diseases 125:406–408

Hirayama T (1981) Non-smoking wives of heavy smokers have a higher risk of lung cancer: a study from Japan. Br Med J 282:183–185

Hirayama T (1984) Cancer mortality in non-smoking women with smoking husbands in a large-scale cohort study in Japan. Preventive Medicine 13:680–690

Humble CG, Samet JM, Pathak DR (1987) Marriage to a smoker and lung cancer risk. Am J Public Health 77:598–602

Jarvis M, Tunstall-Pedoe H, Feyerabend C, Vesey C, Saloojee Y (1984) Biochemical markers of smoke absorption and self-reported exposure to passive smoking. J Epidemiology and Community Health 38:335–339

Kabat GC, Wynder EL (1984) Lung cancer in nonsmokers. Cancer 53:1214–1221

Koo LC, Ho JH-C, Saw D (1984) Is passive smoking an added risk factor for lung cancer in Chinese women? J Experimental and Clinical Cancer Research 3:277–283

Koo LC, Ho JH-C, Saw D, Ho C-Y (1987) Measurements of passive smoking and estimates of lung cancer risk among non-smoking Chinese females. Int J Cancer 39:162–169

Lee PN (1984) Passive smoking. In: Cumming G, Bonsignore G (eds) Smoking and the lung. Plenum Publishing Corporation, p 216

Lee PN (1987) Passive smoking and lung cancer association: A result of bias? Human Toxicology 6

Lee PN (1988a) Difficulties in studying weak associations epidemiologically with particular reference to environmental tobacco smoke exposure and lung cancer (submitted for publication)

Lee PN (1988b) Passive smoking and lung cancer. Association on artefact due to misclassification of smoking habits. Springer (in press)

Lee PN, Chamberlain J, Alderson MR (1986) Relationship of passive smoking to risk of lung cancer and other smoking-associated diseases. Br J Cancer 54:97–105

Matsukura S, Taminato T, Kitano Y, et al (1984) Effects of environmental tobacco smoke on urinary cotinine excretion in nonsmokers. N Engl J Med 311:828–832

Pershagen T, Hrubec Z, Svensson C (1987) Passive smoking and lung cancer in Swedish women. Am J Epidemiology 125:17–24

Pittenger DJ (1985) Passive smoking. N Engl J Med 312:720

Robins J (1986) Risk assessments - exposure to environmental tobacco smoke and lung cancer. Appendix D to environmental tobacco smoke. Measuring exposure and assessing health effects. Committee on passive smoking. Board on environmental studies and toxicology, pp 294–336, Washington, DC. National Research Council. National Academy Press

Trichopoulos D, Kalandidi A, Sparros L, Macmahon B (1981) Lung cancer and passive smoking. Int J Cancer 27:1–4

Trichopoulos D, Kalandidi A, Sparros L (1983) Lung cancer and passive smoking: Conclusion of Greek study. Lancet 2:677–678

Wald NJ, Nanchahal K, Thompson SG, Cuckle HS (1986) Does breathing other people's tobacco smoke cause lung cancer? Br Med J 293:1217–1222

Wells AJ (1986) Misclassification as a factor in passive smoking risk. Lancet 2:638

Wu AH, Henderson BE, Pike MC, Yu MC. Smoking and other risk factors for lung cancer in women. J National Cancer Institute 74:747–751

An Introduction to the Study of Smoking Using Urinary Hydroxyproline

H. Kasuga

Summary

The indicators used to identify the effects of environmental tobacco smoke (ETS) and exposure to low level nitrogen dioxide (NO_2) on the respiratory system are so weak that they can not be detected using traditional markers such as the increased prevalence of respiratory symptoms and a decrease in lung function. After studying urinary hydroxyproline (HOP) starting in 1977, we first reported the significant relationship between HOP and smoking, ETS and NO_2 in the air, in 1981. Since then, the coherent association between urinary HOP and the established pathological and biological development of lung diseases has been studied. Bias problems based on confounding factors, misclassification of nonsmokers and over- or underestimation of ETS effects also have been discussed.

Among articles presented by us during the past 4 years, several papers and some arguments for and against this study on urinary HOP were introduced:

1) The effect of cessation from smoking on the urinary excretion of hydroxyproline [19].
2) A prospective repeated cross-sectional study on the possible health effects caused by automobile exhaust and passive smoking [20].
3) Impact of smoking on the concentration and activity of alpha-1-antitripsin in serum, in relation to the urinary excretion of hydroxyproline. Matsuki H, Kasuga H et al., 1988.
4) Behavior of urinary hydroxyproline and effect of cigarette smoking in silicosis. Osaka F, Kasuga H, Matsuki H et al., 1985.
5) Opinions contrary to the relationship between urinary hydroxyproline and smoking, ETS and NO_2.

Introduction

Hydroxyproline (HOP) is one of the essential constituents in collagen and elastine and is an unique one which is not found in other tissues. Therefore, urinary HOP is regarded to be a potential candidate for the study of the breakdown of lung tissue due to smoking and environmental tobacco smoke (ETS).

As is generally known, the index symptom such as "a persistent cough and phlegm" based on the BMRC Questionnaire [1] is used frequently as a clinical marker, but it is not applicable for ETS effects because prevailing concentrations of ETS are estimated to be less than 1% of an undiluted mixture of sidestream and second-hand mainstream smoke. Therefore, urinary HOP as a biochemical marker for ETS effects appeared on the stage, and a causal relationship between smoking including ETS and its health effect was

We would be glad to apply our technique to more detailed data if we can get them from Hirayama, for instance in order to adjust by occupation of men and age of women, or by occupation of men and by age of women married to a husband who is still alive. We are ready to modify our view if such data can support the alternative hypothesis better than the published data. We do hope, that our calculations give rise to a fruitful discussion. The methods we used here might be of interest to the analysis of other cohort and case control studies.

References

1. Bishop YV, Fienberg StE, Holland PW (1980) Discrete multivariate analysis: theory and practice. MIT Press, Cambridge, pp 97
2. Garfinkel L (1984) Time trends in lung cancer mortality among nonsmokers and a note on passive smoking. J Natl Canc Inst (USA) 66:1061–1066
3. Hartung J, Elpelt B, Klösener KH (1985) Statistik. Lehr- und Handbuch der angewandten Statistik. München Wien Oldenburg, p 501–503
4. Hirayama T (1984) Lung cancer in Japan: effects of nutrition and passive smoking. In: Mizell M, Correa P (eds) Lung cancer: causes and prevention. Chemie, Weinheim, pp 175–195
5. Lee PN, Chamberlain J, Alderson HR (1986) Relationship of passive smoking to risk of lung cancer and other smoking-associated diseases. Br J Cancer (54):97–105
6. Überla K (1987) Lung cancer from passive smoking: hypothesis or convincing evidence? Int Arch Occup Environ Health (59):421–431

2023381520

What Is the Epidemiologic Evidence for a Passive Smoking-Lung Cancer Association?

N. Mantel

Summary

Two survey articles of reports on the association of passive smoking with lung cancer have recently appeared, and also a comprehensive report on the subject of environmental tobacco smoke by a committee of the National Research Council of the United States. The observed excess over a relative risk of unity cannot be explained by chance. Nor can it be fully accounted for by a particular source of bias, the false claims of being non-smokers by individuals who were active or ex-smokers. That possible source of bias leads, in one summary survey, to reducing a relative risk of 1.35 to 1.30, but from 1.34 to 1.15 in the National Research Council report. The latter report suggests that statistical significance would no longer obtain, perhaps, particularly, because of other possible biases. However, to get an estimate of the correct relative risk due to passive smoking, allowance has to be made for actual exposure to passive smoking of those not exposed at home. Thus, the 1.30 is adjusted upwards, by 18 in one survey, to 1.53, but by only 8% in the National Research Council report to 1.24. The National Research Council report had given an anticipated relative risk of 1.1 based on dosimetric considerations. But it is suggested here that that could be as low as 1.05, too low to be detected in an epidemiologic investigation - in any case it would be based on hypothetical assumptions.

In November of 1986 there were two near-simultaneous review articles addressing the subject of passive smoking and lung cancer. One was an invited guest editorial by Blot and Fraumeni in the Journal of the National Cancer Institute, the other a contemporary theme discussion by Wald et al. in the British Medical Journal [1, 2]. There was substantial overlapping in the two articles of the various publications on the subject, and on the basis of which the conclusion of a significant positive association was made. The article by Wald et al. gave, perhaps, more statistical detail about the results of the several studies covered. But, to my mind, there was uncritical acceptance of the results of all the studies. Blot and Fraumeni did suggest that there were some flaws in a particular study, that by Hirayama [3], but decided that any inherent biases in that investigation could not have given rise to the observed elevated risk.

From their overall evaluation of 10 case-control studies (all 10 gave results for females, five separately for males as well) and three prospective studies (two of these covered males separately), which provided 20 separate relative risk (actually odds ratio) values, Wald et al. came up with a summary relative risk of lung cancer due to passive smoking of 1.35 (95% limits 1.19 to 1.54). They trim this down to 1.30 on the basis that some of the presumed non-smokers exposed to passive smoking were actually smokers. Then, on the added basis that even those unexposed to passive smoking at home may still have been exposed when away from home, they raise their estimate of relative risk to 1.53. But note that this last modification presupposes the answer, that passive smoking does